MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2022/2023

DBS5114 – BUSINESS STATISTICS

(All sections / Groups)

03 JULY 2023
9.00 a.m – 11.00 a.m
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This question paper consists of **TWELVE (12)** printed pages with:
   Section A: Ten (10) multiple choice questions (20%)
   Section B: Three (3) structured questions (80%)

2. Answer **ALL** questions.

3. Answer **Section A** in the multiple-choice answer sheet provided and **Section B** in the answer booklet provided.

4. Statistical tables are attached at the end of the question paper.

5. Students are allowed to use non-programmable scientific calculators with no restrictions.

SECTION A: MULTIPLE CHOICE QUESTIONS (20 MARKS)

There are TEN (10) questions in this section. Answer ALL questions on the multiple-choice answer sheet.

1. A population in statistics means a collection of all

   A. men and women
   B. subjects or objects of interest
   C. people living in a country
   D. all subjects or objects in the world

2. Which of the following examples correctly represents quantitative and qualitative variables?

   A. Students' GPAs and number of FOM students
   B. Marital status and salaries of football players
   C. Number of viewers for TV programs and viewer ratings for the TV program
   D. None of the examples above are right

3. A summary measure that is computed to describe a characteristic of an entire population is called

   A. a parameter
   B. a census
   C. a statistic
   D. the scientific method

4. Which of the following statements are true?
   I. All variables can be classified as quantitative or qualitative variables
   II. Qualitative variables can be continuous variables
   III. Quantitative variables can be discrete variables

   A. I and II.
   B. II and III.
   C. II only.
   D. I and III.

5. For a standard normal distribution, what is the probability that z is greater than 1.75?

   A. 0.0401
   B. 0.0459
   C. 0.4599
   D. 0.9599

Continued...

6. The hourly wages of a sample of 130 systems analysts are given below.

   mean = 70 range = 20
   mode = 73 variance = 324
   median = 74

   The coefficient of variation equals

   A. 0.30%
   B. 30%
   C. 5.4%
   D. 25.71%

7. An outlier is _____.

   A. a measurement that is separated from most of the other measurements in the data
   B. a method to predict future observations based on sample data
   C. equivalent to the 90^{th} percentile of a data set
   D. a measure of central tendency

8. If A and B are independent events with $P(A)=0.65$ and $P(A \cap B)=0.26$, then $P(B)=$

   A. 0.400
   B. 0.169
   C. 0.390
   D. 0.650

9. Suppose $P(A)=0.35$. The probability of complement of A is:

   A. 0.35
   B. 0.50
   C. 0.65
   D. -0.35

10. If two events are mutually exclusive, what is the probability that both occur at the same time?

    A. 0
    B. 0.50
    C. 1.00
    D. Cannot be determined from the information given

Continued...

SECTION B: STRUCTURED QUESTIONS (80 MARKS)

There are THREE questions in this section. Candidates MUST answer ALL THREE questions.

QUESTION 1 [25 MARKS]

a) A sample of the number of delivery orders received per day before and during Movement Control Oder (MCO) from a delivery service provider is as follows:

Before	99	101	80	81	92	84	88	83	81	78
During	105	115	125	102	99	135	122	119	110	101

i. Determine the mean and median for the number of delivery orders received per day before and during MCO.

[6 marks]

ii. Determine the variance and coefficient of variation for the number of delivery orders received per day before and during MCO.

[7 marks]

b) A sample of 30 delivery orders from a delivery service provider gives a mean delivery time of 28 minutes with a standard deviation of 8 minutes. Find the 95% confidence interval for the mean delivery time of the delivery service provider.

[5 marks]

c) The amount of time for assembly line workers to complete a product is normally distributed with a mean equal to 15 minutes and a standard deviation of 2 minutes.

i. Calculated a probability that the amount of times to complete a product was smaller than 17 minutes.

[3 marks]

ii. Find a probability that the amount of times to complete a product was greater than 20 minutes.

[4 marks]

Continued...

QUESTION 2 [25 MARKS]

a) After conducting an analysis, a statistics instructor observes significant variation in the number of typographical errors across different books in the new edition. Based on this analysis, the instructor concludes that the distribution of errors has a mean of 1.5 per 100 pages.

i. Find the probability that there are between four to six typographical errors for a textbook with 100 pages.

[4 marks]

ii. Find the standard deviation of this distribution.

[2 marks]

b) Bestflix is a streaming service that allows its customers to watch a wide variety of television shows and movies. A researcher wants to investigate on the number of Bestflix subscribers per household and produced the following probability distribution:

Number of Bestflix subscribers per household	0	1	2	3	4
Probability	0.22	0.2	0.31	0.16	0.11

i. What is the probability of having between two and four Bestflix subscribers per household??

[3 marks]

ii. Determine the average number of Bestflix subscribers per household.

[3 marks]

iii. Calculate the variance and standard deviation for the probability distribution of Bestflix subscribers per household.

[5 marks]

Continued...

c) After the midterm test, the representative council of a Big-Ten university conducted a survey involving a sample of 200 students. The survey aimed to gather information about their level of preparedness prior to the test, specifically in terms of whether they spent time watching movies at the cinema or studying diligently. Additionally, the council also collected data on the students' midterm test results. The findings of the survey are presented below:

The students' method in preparing the midterm test	**The midterm results**	
	Good	***Poor***
Studying hard	80	20
Watching movies at cinema	30	70

i. What is the probability of a student achieving a good result in their midterm test?

[2 marks]

ii. Determine the probability that a student studying hard and get good results.

[3 marks]

iii. Calculate the probability that a student watching movies and get poor results.

[3 marks]

Question 3 (30 Marks)

a) A pharmaceutical company engaged in the manufacture of a newly developed drug. The company claimed that the drug is 80% effective in relieving cold for a period of 15 hours. A sample of 200 patients who suffered from cold were given this drug. It was found that the drug provided relief to 150 patients for at least 12 hours.

i. Determine the sample proportion of patients who were relieved from cold for at least 12 hours.

[3 marks]

ii. Test the company's claim at 5% significance level. Do you think that the company's claim is justified?

[12 marks]

Continued...

b) A scientist wanted to analyze the relationship between the speed of a new electric train (x) and its energy consumption (y). An experimental train is operated at several different speeds and for each speed the energy consumption is measured. These data are shown below:

Speed	45	55	65	70	80	85	90
Energy Consumption	40	39	37	33	30	27	25

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9747
R Square	0.9501
Adj. R Square	0.9401
Standard Error	1.4481
Observations	7

ANOVA

	df	*SS*	*MS*	*F*	*Sig. F*
Regression	1	199.5156	199.5156	95.1490	0.0002
Residual/Error	5	10.4844	2.0969		
Total	6	210			

	Coefficients	*Standard Error*	*t Stat*	*p-value*
Intercept	57.7188	2.5925	22.2634	0.0000
Speed	–0.3531	0.0362	–9.7544	0.0002

i. Determine the correlation coefficient and interpret its meaning.

[3 marks]

ii. State the linear regression equation.

[2 marks]

iii. Interpret the regression coefficients.

[4 marks]

iv. Determine the coefficient of determination and interpret its meaning.

[3 marks]

v. Predict the energy consumption if the speed is 200. Is this prediction reliable?

[3 marks]

End of Questions.

A. DESCRIPTIVE STATISTICS

Mean = $\frac{\sum X_i}{n}$

Standard Deviation (s) = $\sqrt{\frac{\sum X^2}{n-1} - \frac{(\sum X)^2}{n(n-1)}}$

Pearson's Coefficient of Skewness (S_k) = $\frac{3(\overline{X} - \text{Median})}{s}$ or $\frac{\overline{X} - \text{Mode}}{s}$

B. PROBABILITY

P (A or B) = P (A) + P (B) - P (A and B)

P (A and B) = P (A) P (B) if A and B are independent

P (A | B) = P (A and B) / P (B)

Poisson Probability Distribution

If X follows a Poisson Distribution P (λ) where $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$

then the mean = E(X) = λ and variance = VAR(X) = λ

Binomial Probability Distribution

If X follows a Binomial Distribution B(n, p) where $P(X = x) = {}^nC_x p^x q^{n-x}$

then the mean = E(X) = n p and variance = VAR(X) = npq where q = 1 - p

Normal Distribution

If X follows a Normal distribution N(μ , σ) where E(X) = μ and VAR(X) = σ^2

then $z = \frac{X - \mu}{\sigma}$

C. EXPECTATION AND VARIANCE OPERATORS

E (X) = $\sum$[X • P(X)]

VAR (X) = E(X^2) - [E(X)] ^{2}

If E (X) = μ then E (kX) = $k\,\mu$, E(X + Y) = E(X) + E(Y)

If VAR (X) = σ^2 then VAR (kX) = $k^2\,\sigma^2$,

VAR (aX + bY) = a^2VAR(X) + b^2VAR(Y) + 2ab COV(X, Y)

where COV(X, Y) = E(X Y) – [E(X) E(Y)]

Continued...

D. CONFIDENCE INTERVAL ESTIMATION AND SAMPLE SIZE DETERMINATION

(100 - α) % Confidence Interval for Population Mean (σ Known) = $\overline{X} \pm Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$

(100 - α)% Confidence Interval for Population Mean (σ Unknown) = $\overline{X} \pm t_{\alpha/2,n-1}\left(\frac{s}{\sqrt{n}}\right)$

(100 - α)% Confidence Interval for Population Proportion = $p \pm Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$

Sample Size Determination for Population Mean $= n \geq \frac{(Z_{\alpha/2})^2 \sigma^2}{E^2}$

Sample Size Determination for Population Proportion $= n \geq \frac{(Z_{\alpha/2})^2 p(1-p)}{E^2}$

Where E = Limit of Error in Estimation

E. HYPOTHESIS TESTING

One Sample Mean Test	
Standard Deviation (σ) Known	**Standard Deviation (σ) Not Known**
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$	$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$
One Sample Proportion Test	
$Z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$	

Continued...

Two Sample Mean Test

Standard Deviation (σ) Known

$$z = \frac{(\overline{x_1} - \overline{x_2})}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

Standard Deviation (σ) Not Known

$$t = \frac{(\overline{x_1} - \overline{x_2})}{\sqrt{S_p^2\left(1/n_1 + 1/n_2\right)}} \quad \text{where} \quad S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

Two Sample Proportion Test

$$z = \frac{(p_1 - p_2)}{\sqrt{p(1-p)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} \quad \text{where } p = \frac{(n_1 p_1) + (n_2 p_2)}{n_1 + n_2} = \frac{X_1 + X_2}{n_1 + n_2}$$

where X_1 and X_2 are the number of successes from each population

F. REGRESSION ANALYSIS

SIMPLE LINEAR REGRESSION:

Correlation Coefficient

$$r = \frac{\sum XY - \left[\frac{\sum X \sum Y}{n}\right]}{\sqrt{\left[\sum X^2 - \left((\sum X)^2/n\right)\right]\left[\sum Y^2 - \left((\sum Y)^2/n\right)\right]}} = \frac{COV(X,Y)}{\sigma_X \sigma_Y}$$

Regression Coefficient

$$b_1 = \frac{\sum XY - \left[\frac{\sum X \sum Y}{n}\right]}{\left[\sum X^2 - \left((\sum X)^2/n\right)\right]}, \qquad b_0 = \overline{Y} - b_1 \overline{X}$$

Continued...

MULTIPLE LINEAR REGRESSION:

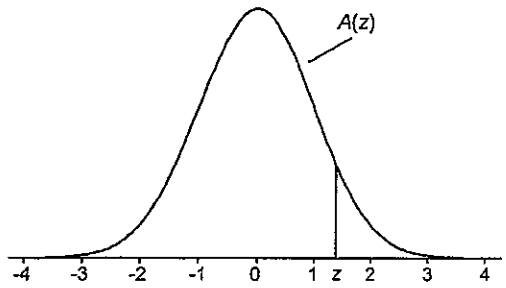
Adjusted r-square $= 1 - \left[\frac{(1-r^2)(n-1)}{(n-p-1)}\right]$ where p = number of independent variables

ANOVA Table for Multiple Linear Regression			
Source	**Degrees of Freedom**	**Sum of Squares**	**Mean Squares**
Regression	p	SSR	MSR = SSR/p
Error	$n-p-1$	SSE	MSE = SSE/($n-p-1$)
Total	$n-1$	SST	
Test Statistic for Significance of the Overall Regression Model = F = MSR/MSE			
Test Statistic for Significance of each Explanatory Variable = $t^* = b_i / S_{bi}$ and the Critical $t = t_{(n-p-1),\alpha/2}$			

Continued...

TABLE A.1

Cumulative Standardized Normal Distribution



$A(z)$ is the integral of the standardized normal distribution from $-\infty$ to z (in other words, the area under the curve to the left of z). It gives the probability of a normal random variable not being more than z standard deviations above its mean. Values of z of particular importance:

z	$A(z)$	
1.645	0.9500	Lower limit of right 5% tail
1.960	0.9750	Lower limit of right 2.5% tail
2.326	0.9900	Lower limit of right 1% tail
2.576	0.9950	Lower limit of right 0.5% tail
3.090	0.9990	Lower limit of right 0.1% tail
3.291	0.9995	Lower limit of right 0.05% tail

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999							

Continued...

t-Distribution Critical Values

Table entry for p and C is the value of t^* with probability p lying above it and probability C lying between $-t^*$ and t^*.

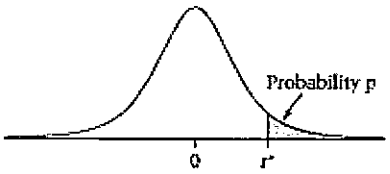


TABLE B *t*-Distribution Critical Values

	Tail Probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

End of Page.